

Globus.org Tutorial

Lisa Childers

Paul Davé

GlobusWorld 2010



Welcome To Tutorial Day!

Globus.org Tutorial: 9am-noon

Lunch: noon-1pm

Globus Toolkit 5.0 Tutorial: 1-4pm

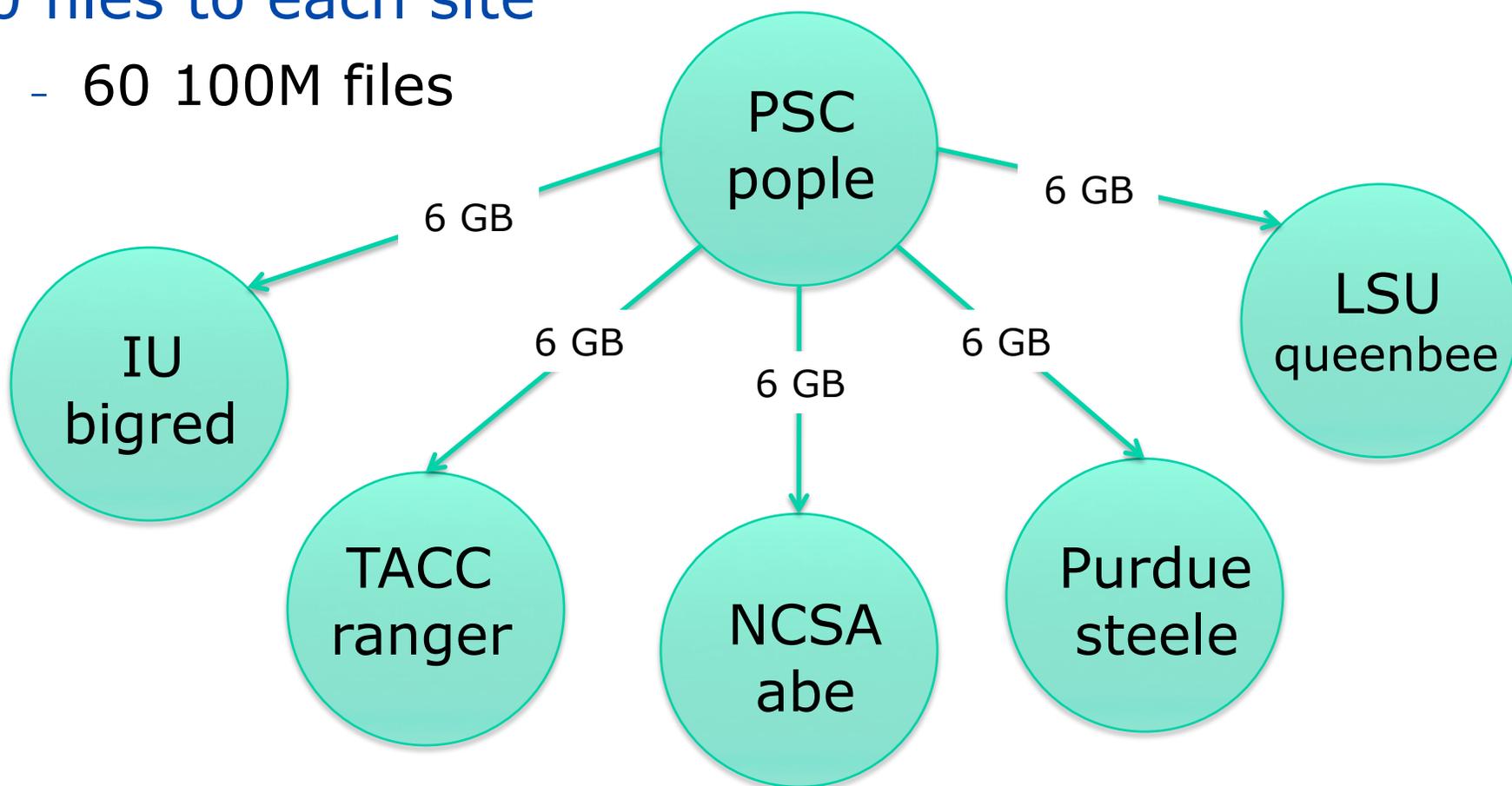
To begin, I will fire up some transfers...



300 Files, 30 GB

60 files to each site

- 60 100M files

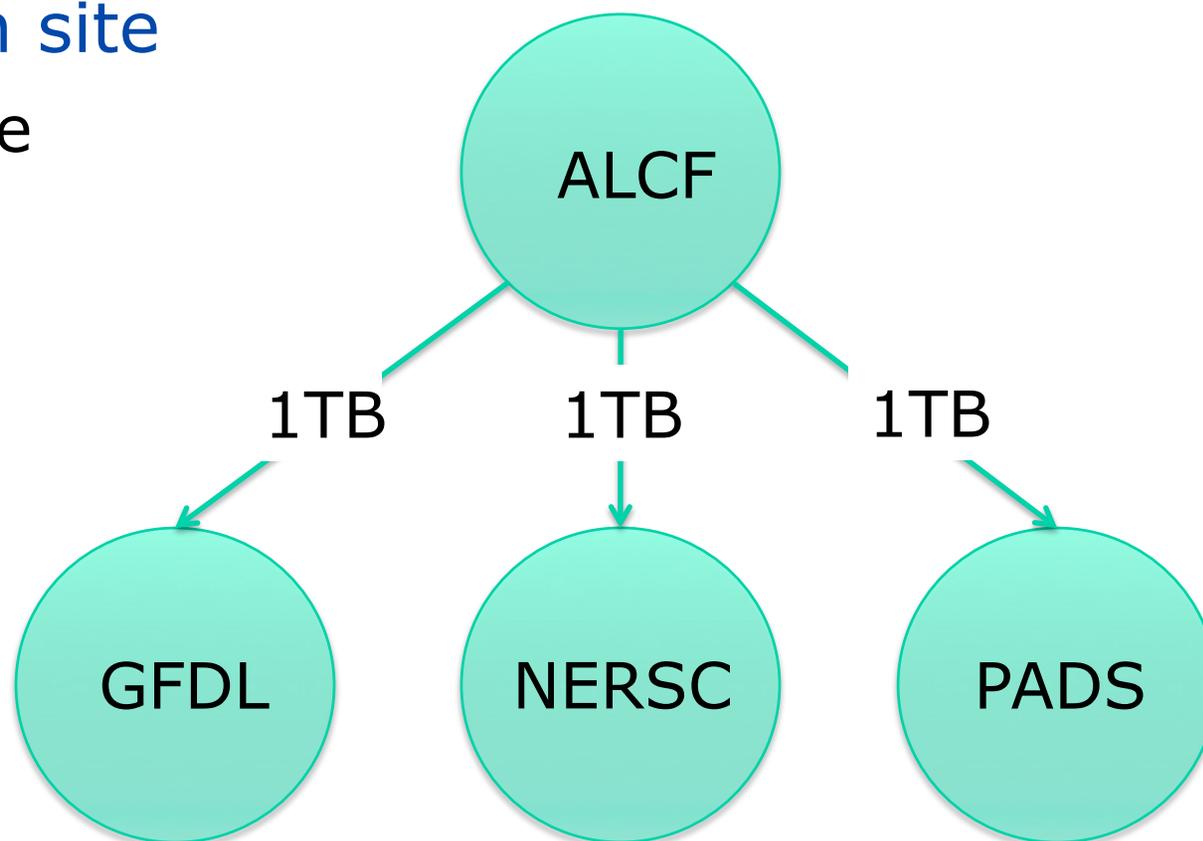




24 Files, 3TB

8 files to each site

- 1 498G file
- 2 200G
- 1 100G
- 1 20G
- 1 10G
- 1 5G
- 1 1G





the globus alliance

www.globus.org

Globus.org Tutorial Outline

- **9:00 am - Introduction**
 - What problem are we trying to solve?
 - What is our approach for solving the problem?
- 10:00 am - Break
- 10:30 am - Current Status and Plans
 - What does Globus.org look like today?
 - What might it look like tomorrow?
- 11:30 am – Globus.org User Services
- Wrap-up
 - Present transfer results
 - Acknowledgements

Key Problems and Requirements

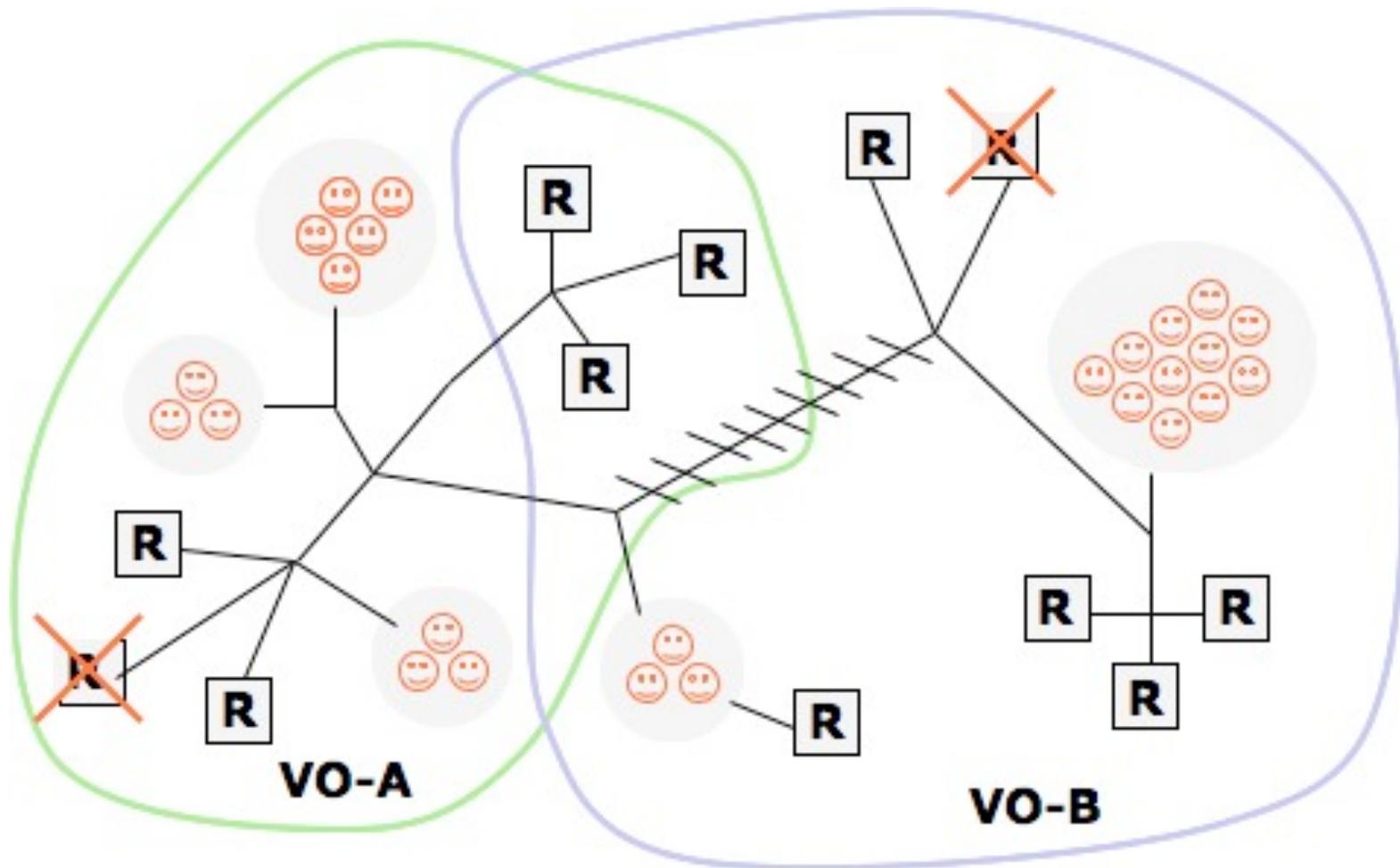
**Problem #1:
Facilitate cross-administrative domain
interactions while at the same time
protecting local autonomy**



the globus alliance

www.globus.org

Facilitate The Work Of Virtual Organizations





the globus alliance

www.globus.org

Support Heterogeneity and Local Control

- Local sites have their own
 - User policies
 - Authorization mechanisms
 - Data privacy policies
 - Hardware
 - Software stacks
 - Service and network configurations
- The sites should be able to share their resources without losing control over them



Key Requirements

- Globus.org should strive to be compatible with the resource owner's preferred software stack
 - Avoid imposing new software requirements
 - Support existing security mechanisms

Problem #2:
**Most users lack the time and
inclination to become experts in
distributed computing technology**

Overview of Reported User Goals

Perspectives on Distributed Computing User Interviews



<http://www.mcs.anl.gov/~childers/perspectives/>



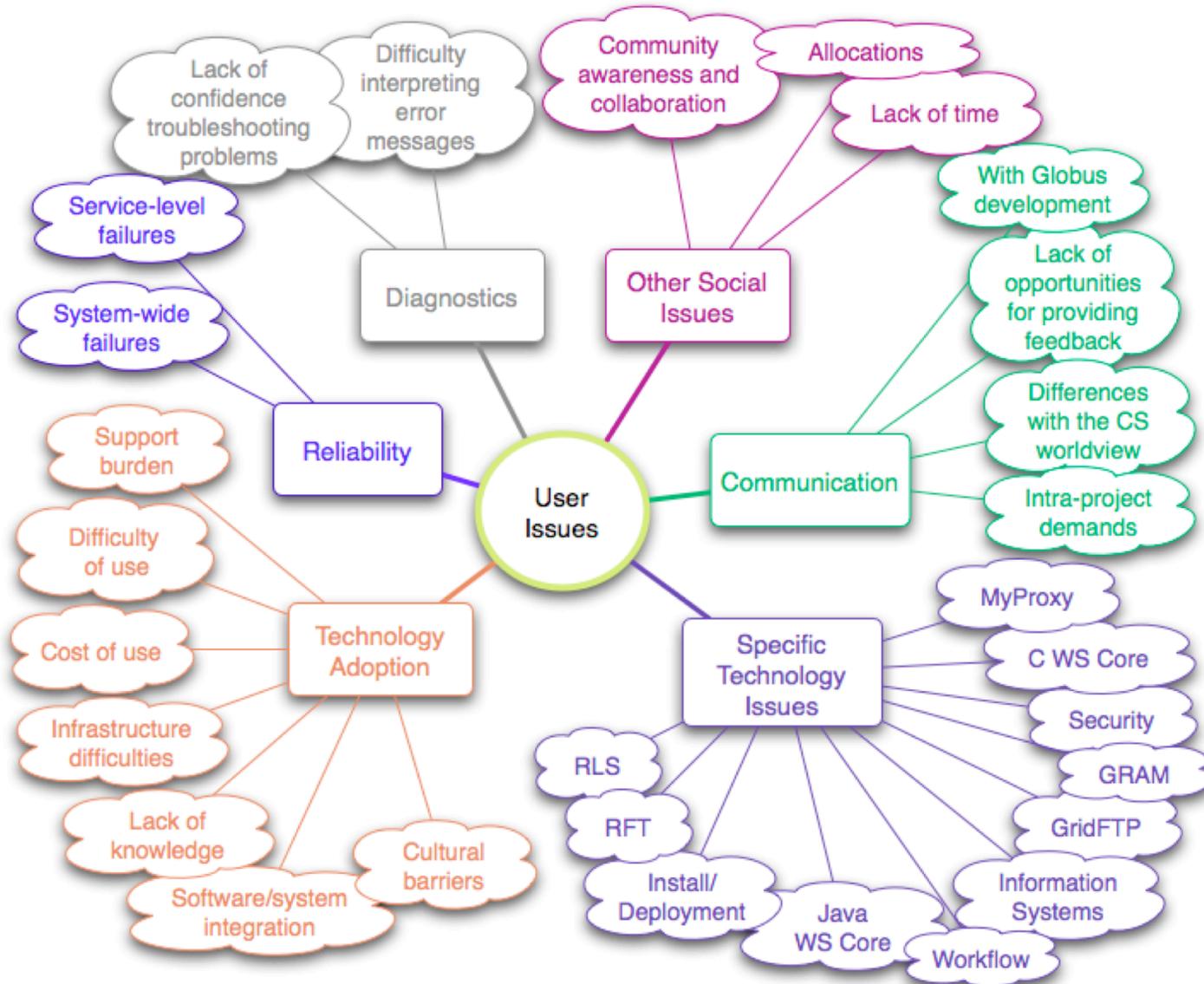
Key Requirements

- Implement familiar user interfaces
 - Technology interactions should require no special expertise
- Minimize end-user software installation requirements
- Ease the infrastructure providers' support burden

**Problem #3:
Both end-users and infrastructure
providers struggle mightily with
wide-area technology failures**

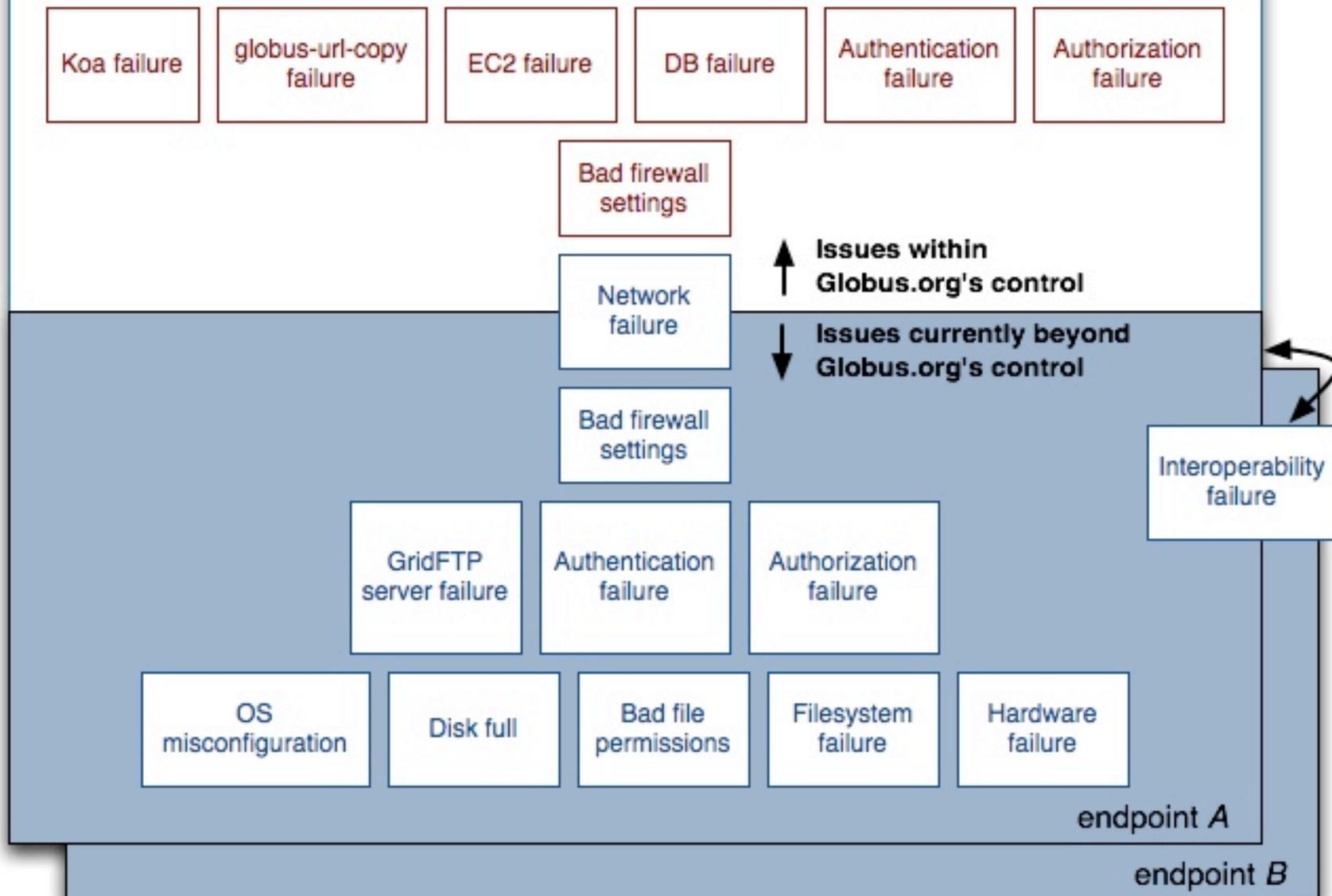
Overview of Reported User Issues

Perspectives on Distributed Computing User Interviews



<http://www.mcs.anl.gov/~childers/perspectives/>

Potential Sources of File Movement Problems





the globus alliance

www.globus.org

Key Requirements

- Manage an increasing number of technology failures on behalf of the user
- Provide users and resource owners with enough information (in words they can understand) to efficiently resolve problems
- Send notifications of interesting events
 - Now: send an email when a transfer completes
 - Someday: Give end-users and resource providers a heads-up about potential problems



the globus alliance

www.globus.org

Problem #4: More Data Is Coming





the globus alliance

www.globus.org

Anticipated ALCF Bandwidth Requirements *

- 0-2 years: 10s of TB/day
- 2-5 years: 100s of TB/day
- 5+ years: PBs/day

* Office of Advanced Scientific Computing Research Network Requirements Workshop, April 15-16, 2009



the globus alliance

www.globus.org

Practical Requirements

- Meet the upcoming CEDPS challenges
 - Just recently met a 100k file, 100MB challenge
 - Start moving 40TB/day of GFDL data next year

What additional problems and requirements do you have?

The Current Globus.org Implementation



Distributed Data Nodes

GridFTP

GridFTP

GridFTP

GridFTP

GridFTP

GridFTP

GridFTP

GridFTP

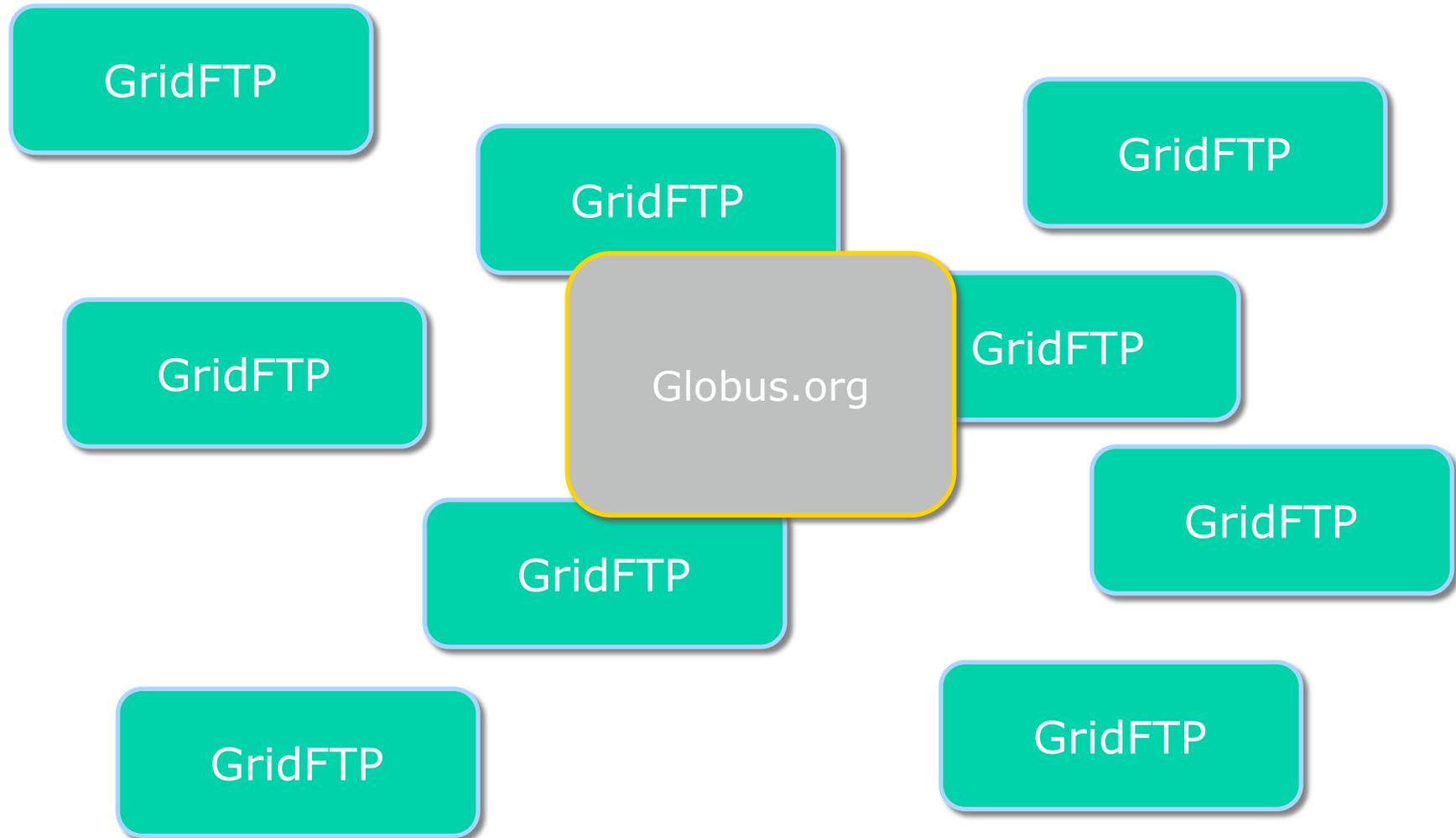
GridFTP



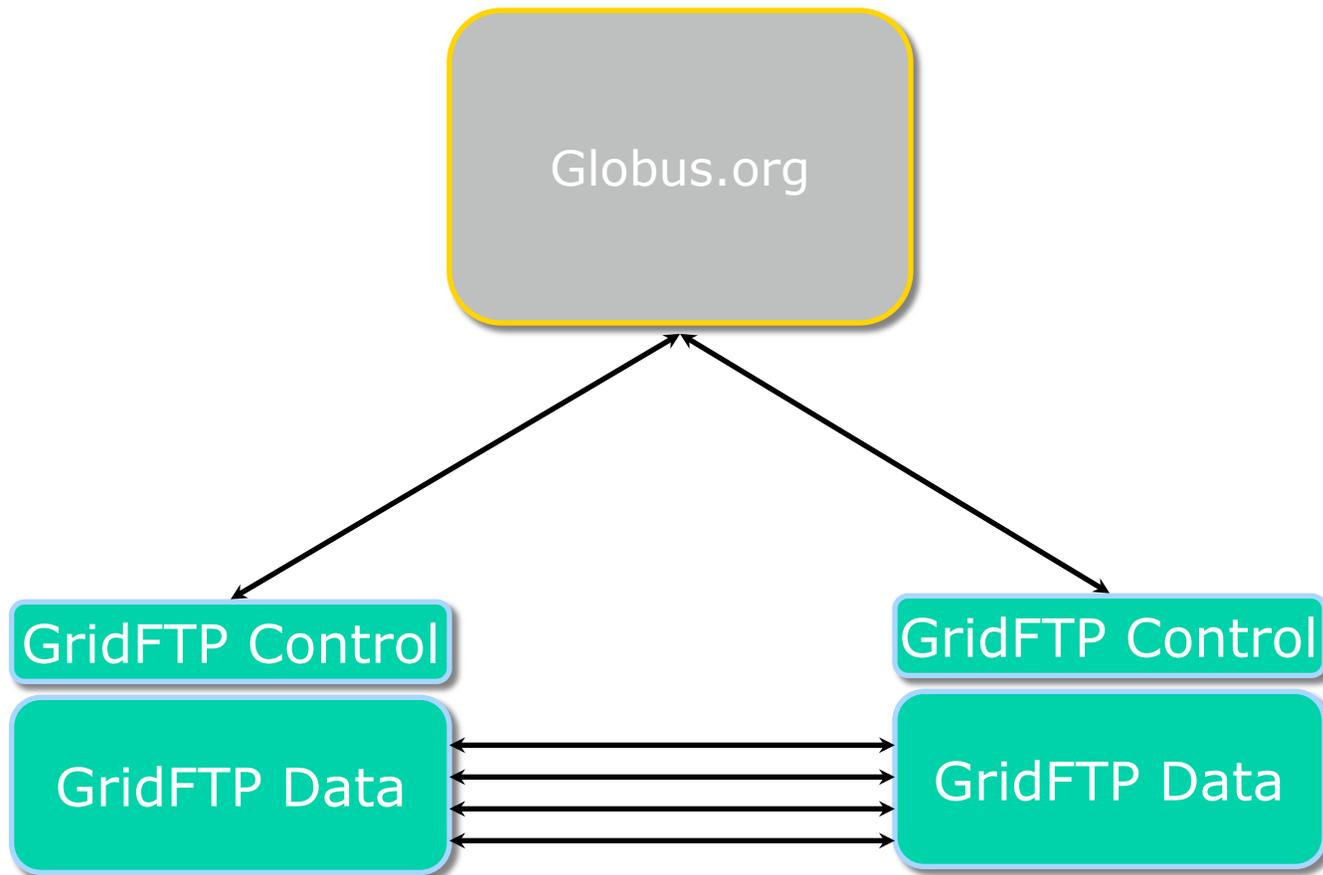
the globus alliance

www.globus.org

Globus.org Operates At The Collective Layer

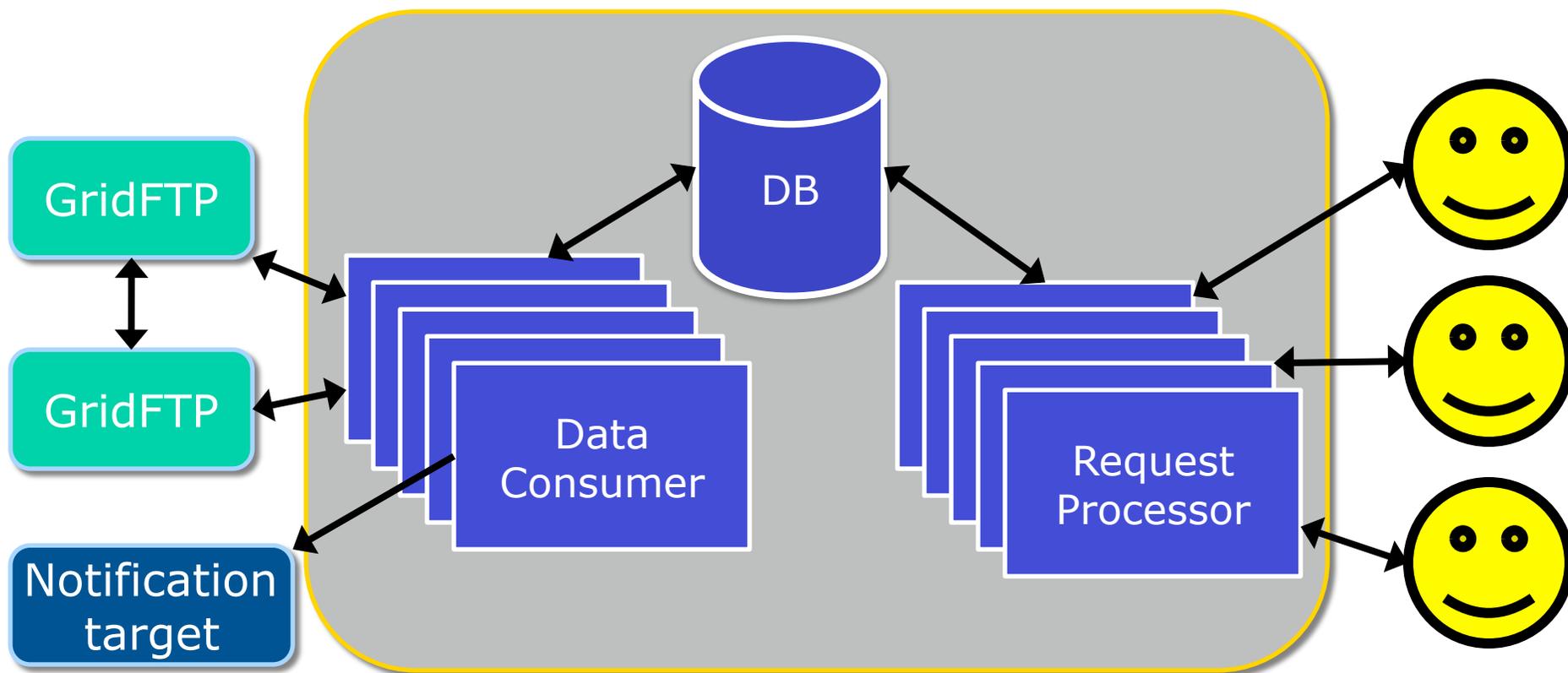


Globus.org Manages 3rd-Party Transfers





A Peek Inside Globus.org





Anatomy of a Globus.org CLI Call

```
ssh -t <user>@<machine> <command> <options> <params>
```

Globus.org command

Globus.org instance

Globus.org userid

Use as needed to hide password text



the globus alliance

www.globus.org

gsissh can be used instead of ssh

```
ssh -t <user>@<machine> <command> <options> <params>
```

```
gsissh -p 2222 -o 'GSSAPITrustDNS no'
```

Override DNS checks because the host certs do not currently match the Amazon IPs (will not be required in future releases)

GSI-OpenSSH server port



A Recap: What is Globus.org?

- The latest iteration of Globus software
 - The same Globus vision, but an updated approach
- Hosted services
 - Data movement initially
 - Execution and other services to follow
- The Globus Toolkit isn't going away
 - Contains tools and services for resource owners
 - Compatible with Globus.org hosted services



Key Goals

- Provide scientists with easy access to advanced computing resources
 - Familiar user interfaces
 - Technology interactions requiring no special expertise
 - No software to install
 - Support for well-known community and international resources
 - Ability to customize working environment
- Enable users to focus on domain-specific work
 - Manage technology failures
 - Notifications of interesting events
 - Provide users with enough information to resolve problems



the globus alliance

www.globus.org

Globus.org Tutorial Outline

- 9:00 am - Introduction
 - What problem are we trying to solve?
 - What is our approach for solving the problem?
- 10:00 am - Break
- **10:30 am - Current Status and Plans**
 - What does Globus.org look like today?
 - What might it look like tomorrow?
- 11:30 am – Globus.org User Services
- Wrap-up
 - Present transfer results
 - Acknowledgements

Demonstration



the globus alliance

www.globus.org

Globus.org Tutorial Outline

- 9:00 am - Introduction
 - What problem are we trying to solve?
 - What is our approach for solving the problem?
- 10:00 am - Break
- 10:30 am - Current Status and Plans
 - What does Globus.org look like today?
 - What might it look like tomorrow?
- **11:30 am – Globus.org User Services**
- Wrap-up
 - Present transfer results
 - Acknowledgements



Topics

- Role of user services
- Overview of engagement model
- User examples
- Advantages of early-user engagement
- How best to proceed



Technology + Services = Capability

- Start with Globus.org as a hosted solution
- Apply a user engagement approach to identify user needs
- Provide technology and services that meet these needs



the globus alliance

www.globus.org

Engagement Model

- Jointly discuss your requirements and objectives – capture your use case
- Review use case and map out steps for meeting your goals
- Create Globus.org accounts
- Provide hands-on guidance, testing and support to achieve success



the globus alliance

www.globus.org

Support Takes on Many Forms

- Globus.org walkthroughs, customized training
- Run sanity tests against targeted use-cases
- Assistance with configuring end-user environment
- Provide priority access to engineering-level support



the globus alliance

www.globus.org

Example Engagements

- Geophysical Fluid Dynamics Laboratory and DOE Leadership Computing Facilities
 - Reviewed Use Case
 - Setup service to demonstrate data movement between ALCF and GFDL
 - Worked through environment issues
 - Demonstrated successful data movement
 - *Validating that the service can meet user needs*
 - *Preparing to hand-off to users*



the globus alliance

www.globus.org

Example Engagements

- **NERSC**
 - Presented service capabilities
 - Setup service for internal evaluation
 - *Gathering feedback*
 - *Working towards enabling Globus.org data movement service for NERSC*



Looking For New Users

- We would like to know more about your data movement needs
- Please take a moment before the lunch break to fill-out a questionnaire
- We will be happy to follow-up and answer additional questions that you might have



the globus alliance

www.globus.org

Advantages in Early User Engagement

- High-levels of user support (via user services and engineering teams)
- You help to drive development priorities
- We can begin addressing your areas of need ASAP



How Best to Proceed

- Let us know that you are interested
 - Fill-out Questionnaire
 - Email support@globus.org
 - Talk to one of the Globus.org team members today!
- We will assign a user liaison as a starting point
- We can then jointly map-out steps to help you to begin using Globus.org



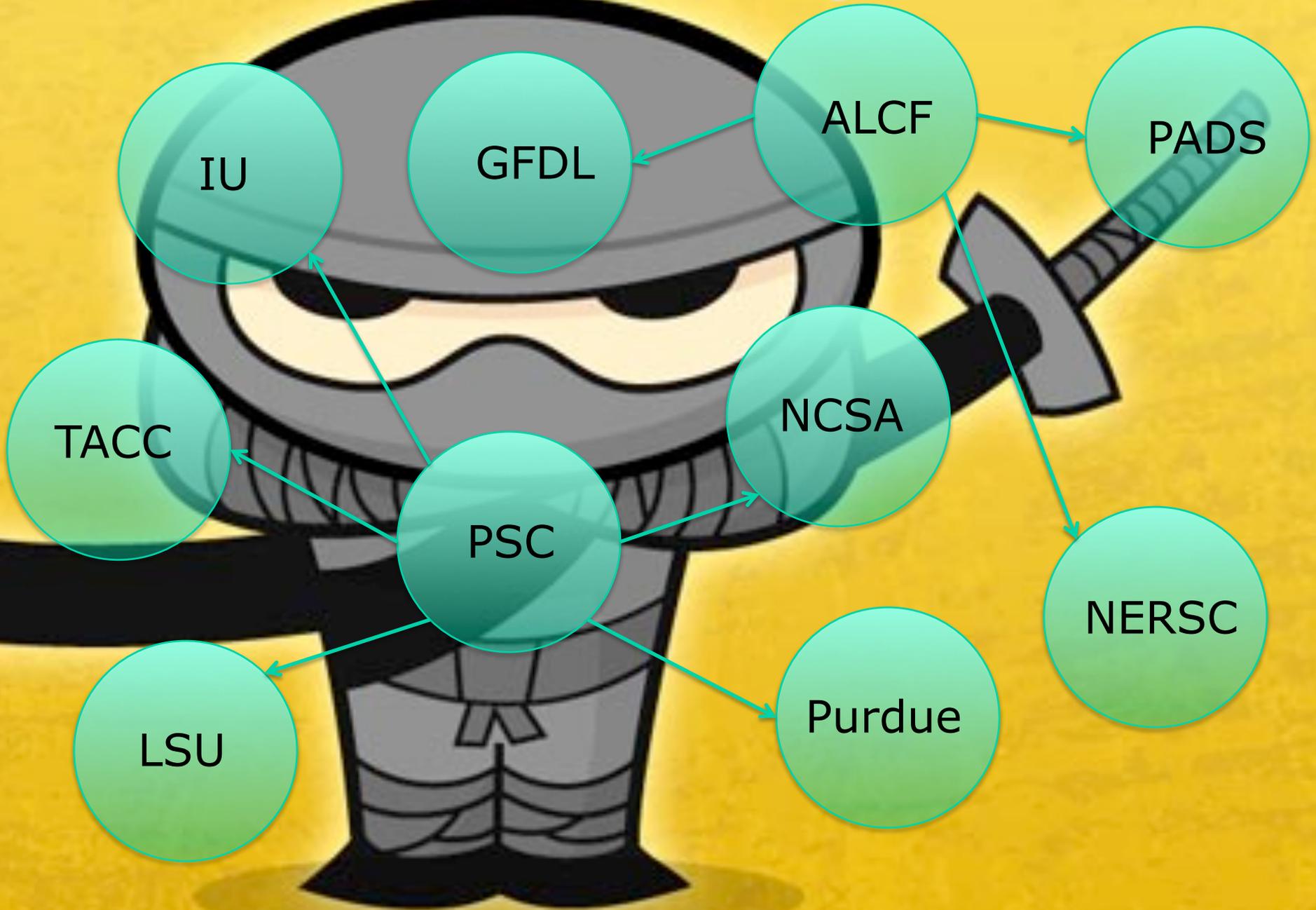
the globus alliance

www.globus.org

Globus.org Tutorial Outline

- 9:00 am - Introduction
 - What problem are we trying to solve?
 - What is our approach for solving the problem?
- 10:00 am - Break
- 10:30 am - Current Status and Plans
 - What does Globus.org look like today?
 - What might it look like tomorrow?
- 11:30 am – Globus.org User Services
- **Wrap-up**
 - Present transfer results
 - Acknowledgements

Globus.org is a "Fire and Forget" Transfer Ninja





Acknowledgements

- **Globus Alliance Colleagues**
 - Especially John Bresnahan, Fred Dech, Ian Foster, Raj Kettimuthu, Jack Kordas, Mike Link, Stu Martin, Bill Mihaló, Mei Hui Su, Steve Tuecke
- **Harvey Wasserman, NERSC User Services**
- **Funders**
 - DOE and NSF

And thanks to you for your
interest and attention!

Welcome To Tutorial Day!

Globus.org Tutorial: 9am-noon

Lunch: noon-1pm

Globus Toolkit 5.0 Tutorial: 1-4pm